

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:47:23 ; Search time 210.42 Seconds
(without alignments)
6.825 Million cell updates/sec

Title: US-09-331-631a-8_COPY_120_161

Perfect score: 245
Sequence: 1 SQRFQECQCHQHQEQRP.....QQCVRECRKXQENPMRGER 42

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_36.*
1: /SIDSL/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqp/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	245	100.0	590	19	W62832
2	111.5	45.5	525	19	W62831
3	111.5	45.5	566	13	R20181
4	99	40.4	625	19	W62830
5	98	40.0	666	19	W62829
6	94	38.4	666	19	W62828
7	74	30.2	28	19	W62841
8	70	28.6	637	19	W62837
9	66	26.9	816	16	R71111
10	66	26.9	816	20	V33494
11	64	26.1	1898	20	V30795
12	63	25.7	71	20	Y09181

13	63	25.7	184	12	R14193	Napin 1-2 gene pro
14	63	25.7	184	15	R54981	Napin protein from
15	63	25.7	184	16	R76694	B. campestris napin
16	63	25.7	184	18	W15415	Napin encoded by c
17	63	25.7	184	19	W30526	Brassica campestris
18	63	25.7	184	21	V49926	Brassica campestris
19	63	25.7	189	9	P81140	Sequence encoded b
20	63	25.7	189	15	R54980	Napin protein from
21	62.5	25.5	35	13	R21079	Antimicrobial maize
22	62.5	25.5	175	14	R33390	Albumin-type prote
23	62	25.3	86	20	W95073	GST-HD fusion prot
24	62	25.3	86	20	W95078	GST-HD fusion prot
25	62	25.3	94	20	W95075	GST-HD fusion prot
26	62	25.3	94	20	W95080	Amino acid sequenc
27	62	25.3	108	20	W95071	Amino acid sequenc
28	62	25.3	108	20	W95076	Amino acid sequenc
29	62	25.3	184	19	W30670	Napin seed storage
30	62	25.3	910	20	V22191	Mouse brain CNG-I
31	61.5	25.1	33	19	W62816	Zea mays antimicro
32	61.5	25.1	593	19	W62816	Zea mays antimicro
33	61.5	25.1	1211	19	W47028	Human N-proteinase
34	61.5	25.1	425	12	R13792	E75B exon B1 poly
35	61	24.9	678	14	R42087	D. melanogaster do
36	60.5	24.7	409	20	W90342	G. max SBP2 protei
37	60.5	24.7	489	20	W90341	Part of the sequen
38	60	24.5	140	10	P91891	Polypeptide tagge
39	59	24.1	371	20	W73369	Mouse SRV-related
40	58	23.7	395	17	W03474	Deltex protein. H
41	58	23.7	737	16	R76640	Deltex protein. H
42	58	23.7	737	16	R76639	Drosophila Deltex
43	58	23.7	737	18	W18317	Amino acid sequenc
44	58	23.7	2074	21	V54319	Mouse pancreatic
45	57	23.3	214	20	Y06303	

ALIGNMENTS

RESULT 1	
ID W62832	standard; Protein: 590 AA.
XX AC W62832:	
XX DT 27-OCT-1998	(first entry)
XX DE Gossypium hirsutum antimicrobial protein.	
XX KW antimicrobial protein; Infestation: control.	
XX OS Gossypium hirsutum.	
XX PN W09827805-A1.	
XX PD 02-JUL-1998.	
XX PE 22-DEC-1997;	97WO-AU00874.
XX PF 20-DEC-1996;	96AU-0004275.
XX PR (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
XX PA Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;	
XX PI WPI, 1998-377279/32.	
XX DR Novel anti-microbial protein from e.g. Macadamia integrifolia.	
XX PT Useful for controlling microbial infestations of plants or mammals	
XX PS Claim 1; Page 49-51; 96pp; English.	
XX CC The sequence is that of an antimicrobial protein which can	
CC be used to control microbial infestations in plants and mammalian	


```

XX      27-OCT-1998      (first entry)
DT
XX      Stenocarpus sinuatus antimicrobial protein.
DE
XX      antimicrobial protein; infestation; control.
KW
XX      Stenocarpus sinuatus.
OS
XX      WO9827805-A1.
PN
XX      02-JUL-1998.
PD
XX      22-DEC-1997;    97WO-AU00874.
PF
XX      20-DEC-1996;    96AU-0004275.
PR
XX      (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PA
XX      Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
PI
XX      WPI; 1998-377279/32.
DR
XX      Novel anti-microbial protein from e.g. Macadamia integrifolia -
DT
XX      27-OCT-1998      (first entry)
DT
XX      Stenocarpus sinuatus antimicrobial protein.
DE
XX      antimicrobial protein; infestation; control.
KW
XX      Stenocarpus sinuatus.
OS
XX      WO9827805-A1.
PN
XX      02-JUL-1998.
PD
XX      22-DEC-1997;    97WO-AU00874.
PF
XX      20-DEC-1996;    96AU-0004275.
PR
XX      (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PA
XX      Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
PI
XX      WPI; 1998-377279/32.
DR
XX      Novel anti-microbial protein from e.g. Macadamia integrifolia -
DT

```

```

xx useful for controlling microbial infestations of plants or mammals
ps Claim 1; Page 60-62; 96pp; English.
cc The sequence is that of an antimicrobial protein which can
cc be used to control microbial infections in plants and mammalian
cc animals.
xx Sequence      637 AA;
SQ
OY Query Match          28.6%; Score 70; DB 19; Length 637;
   Best Local Similarity 30.6%; Pred. No. 1.6;
Matches 15; Conservative 7; Mismatches 13; Indels 14; Gaps 2
DB    6 QECOOHCQHQBQRPKQQCVRECKEKYEEN-----PARGC 41
     | - | - | : | : |||::||| ::||
Db    42 gqcvqrgrerpr-ysharcvgecrddqqghrhegeeqrgrgwhge 89

RESULT 9
ID R71111 standard; Protein; 816 AA.
XX
AC R71111;
XX
DT 22-SEP-1995 (first entry)
XX
DE Spincocerebellar ataxia type 1 (SCA 1) protein product.
KM Spincocerebellar ataxia type 1; SCA 1; presymptomatic diagnosis.
XX Homo sapiens.
OS
PN WO9501437-A.
PD 12-JAN-1995.
PF 29-JUN-1994; 94WO-US07336.
PR 29-JUN-1993; 93US-0084365.
PP 28-JUN-1994; 94US-0267803.
PX (MINU ) UNIV MINNESOTA.
PY
PS Chung M, Orr HT, Zoghbi HY;
PI WPI; 1995-061001/08.
DR N-Psdb; QB4793.
DZ New autosomal dominant spincocerebellar ataxia type 1 nucleic acid
PT - used to develop prods. for detection or presymptomatic
PT diagnosis of a SCA1 disorder
PS Claim 31; Fig 15; 111pp; English.
XX
CC OB4793 is a new autosomal dominant spincocerebellar ataxia type 1
CC (SCA 1) nucleic acid; it encodes the protein product described in
CC R71111. Both the nucleic acid and the protein can be used to develop
CC products, for the presymptomatic detection of a SCA 1 disorder.
CX
SQ Sequence      816 AA;

Query Match           26.9%; Score 66; DB 16; Length 816;
Best Local Similarity 52.2%; Pred. NO. 5.9;
Matches 12; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
DB    2 GROFEOCOHCHQEORPEKKOO 24
     |:|:||| |||:|:||||

```

	RESULT	10
XX	ID	Y33494 standard; Protein: 816 AA.
AC	YY3494;	
XX	Df	19-JAN-2000 (first entry)
XX	DE	Human SCAl protein.
KM	Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;	
KM	Huntingtin polypeptide; Machado-Joseph disease; SCA1; SCA2; SCA6;	
KM	atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;	
KM	Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;	
KX	dentatorubropallidoluysian atrophy; cell proliferation; cell survival;	
KW	neoplastic; malignant; autoimmune; fibrotic.	
OS	Homo sapiens.	
XX	WO994594-A1.	
PJ	16-SEP-1999.	
PD	11-MAR-1999; 99WO-US05250.	
PF	PR	12-MAR-1998; 98US-0041886.
PA	(BURN-) BURNHAM INST.	
Bredesen DE, Rabizadeh S;		
WP1; 1999-561617/47.		
N-PsDB; Z23427.		
New proapoptotic dependence peptides, used to develop products for treating, e.g., Alzheimer's disease -		
treatment,		
Disclosure; Page 127-130; 199pp; English.		
This invention describes novel pure proapoptotic dependence peptides which comprise a sequence of an active dependence domain selected from dependence polypeptides consisting of p75NTR, androgen receptor, DCC, huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2, SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable of inducing cell death and can be used to develop products to mediate or inhibit apoptosis. The methods can be used for reducing the severity of a proapoptotic dependence domain mediated pathological conditions e.g. Huntingtton's disease, Alzheimer's disease, Kennedy's disease, Spinocerebellar ataxias, dentatorubripallidoluysian atrophy, Machado-Joseph disease, stroke or head trauma. They can also be used for reducing the severity of a pathological condition mediated by upregulated cell proliferation or cell survival e.g. neoplastic, malignant, autoimmune or fibrotic conditions. This sequence represents the human SCAl polypeptide described in the method of the invention.		
SQ Sequence	816 AA:	
OY	2 ORFOECOOHCQHCEORPEKKOQ	24
I:I I: I: IIII I::I: ::::II		
Query Match	Best Local Similarity	Score 66; DB 20; Length 816; Pred. No. 5.9;
Matches 12; Conservative	7; Mismatches	4; Indels 0; Gaps 0.0;
Db	200 gqgqqggghqhghqgqqgqqgqq	222
RESULTS	11	
YID Y30795		
standard; Protein: 1898 AA.		
Y30795;		

[illegible]

```

XX 02-SEP-1997; 97US-0922201.
XX PA (SEQU-) SEQUENOM INC.
XX P1 Higgins GS, Koester H, Little D, Lough D:
DR N-PSDB; X37239.
XX WP1: 1999-312406/26.
XX DR N-PSDB; X37239.
XX PS Mass spectrometry detection of polypeptides useful for determining
PT genetic predisposition to Huntington's disease and prostate cancer
XX Example 1; Page 130; 134pp; English.
XX The invention relates to a process for determining the identity of a
CC target polypeptide, using mass spectrometry to determine the molecular
CC mass, and comparing this with molecular mass of a known polypeptide. The
CC method comprises (a) obtaining the polypeptide by in vitro translation,
CC or transfection then translation, of a nucleic acid encoding the
CC polypeptide; (b) determining the molecular mass of at least one of the
CC obtained fragments by mass spectrometry; and (c) comparing the fragment'
CC molecular mass with that of the fragments of a corresponding known
CC polypeptide. The methods are used to detect the presence or disposition
CC for genetic diseases or conditions particularly those caused by an
CC abnormal number of trinucleotide repeats in a gene, such as Huntington's
CC disease, prostate cancer, Fragile X syndrome type A, myotonic dystrophy
CC type I, Kennedy disease, Machado-Joseph disease, and dentatorubral or
CC pallidolysian atrophy, spinobulbar muscular atrophy or ageing, and also
CC for genotyping, in forensic analysis and parental testing, where the
CC nucleotide repeats, preferably di-, tri-, tetra-, or penta-nucleotide
CC repeats, are quantified.
XX Sequence 71 AA:
SO
Query Match 25.7%; Score 63; DB 20; Length 71;
Best Local Similarity 46.2%; Pred. No. 1.2;
Matches 12; Conservative 7; Mismatches 7; Indels 0; Gaps 0
OY 2 ORQFECQDHCHQEOREPKKOCVVR 27
| : | : ||| ||| : : : | |
Db 17 qqqqqqqqhghqggqqgqqghsr 42
RESULT 13
ID R14193 standard; Protein; 184 AA.
AC R14193;
XX R14193;
DT 11-DEC-1991 (first entry)
DE Napin 1-2 gene product.
XX Desaturase; expression cassette; B. campestris.
XX W09113972-A.
XX PN W09113972-A.
XX PD 19-SEP-1991.
XX PF 14-MAR-1991; 91MO-U001746.
PR 14-NOV-1990; 90US-0615784.
PR 16-MAR-1990; 90US-0494106.
PR 13-AUG-1990; 90US-0567373.
XX (CALG-) CALGENE INC.
PA Thompson G, Knauf V;
PI
XX WP1: 1991-295627/40.
DR N-PSDB; Q13969.
XX
```

```

PT DNA encoding a plant desaturase used for modifying the satd.
PT fatty acid compsn. of plant cells and plant seeds
XX
XX
XX Disclosure: Fig 10; 128pp; English.
PS
XX
XX An expression construct contg. a plant desaturase suitable for plant
CC transformation is produced. 5' upstream and 3' downstream sequences
CC obtainable from the B. campestris napin gene here are used.
CC See also Q13963-69.
XX
SQ Sequence 184 AA;

Query Match 25.7%; Score 63; DB 12; Length 184;
Best Local Similarity 25.4%; Pred. No. 3;
Matches 18; Conservativity 7; Mismatches 8; Indels 38; Gaps 3
QY 2 OKFOE-----COOHCHOO-----EORPEKKQCCV 26
Db 48 rrefgagqhlkscqgwlhkhqamsgsgpswtldgetdfedvengqgpgqfppllqgc- 106
QY 27 RECREKYQENP 37
Db 107 --cnelhgeep 115

RESULT 14
R54981
ID R54981 standard; Protein; 184 AA.
XX
AC R54981;
XX
DT 07-DEC-1994 (first entry)
XX
DE Napin protein from Brassica campestris.
XX
DE enhance seed oil production; acyl carrier protein; ACP-I.
XX
OS Brassica campestris.
XX
PN US5315001-A.
XX
PD 24-MAY-1994.
XX
PF 31-JUL-1986; 86US-0891529.
XX
PR 31-JUL-1986; 86US-0891529.
PR 28-JUL-1987; 87US-0078924.
PR 15-NOV-1989; 89US-0437764.
PR 28-JAN-1992; 92US-0826696.
XX
PA (CALJ ) CALGENE INC.
XX
XX Knauf VC, Kridl JC;
PI WPI, 1994-166646/20.
DR N-PSDB; Q65428.
XX
XX DNA encoding plant acyl carrier protein - used to enhance prodn.
PT of seed oil and to modulate the fatty acid compsn. of the oil
XX
PS Disclosure: Fig 4E-4G; 28pp; English.
XX
XX R54981 shows a Brassica napin protein, generated from lambda CGN1-2
CC clone. The napin gene promoter is regulated during seed maturation.
CC This can be used to control the expression of an ACP in seed to
CC enhance the production of seed oil and modulate the fatty acid
CC composition. The DNA can also be used to prepare probes.
XX
SQ Sequence 184 AA;

Query Match 25.7%; Score 63; DB 15; Length 184;

```


